**International Conference on Applied Energy 2019**
**Aug 12-15, 2019, Västerås, Sweden**
**Paper ID: 568**

# AUTOMATED PARAMETER IDENTIFICATION: ROBUST MODEL VALIDATION OF A COMPRESSION CHILLER BASED ON AN UNCERTAINTY AND CONSISTENCY ANALYSIS

Oliver Doelle[1,2], Arvid Amthor[1], Christoph Ament[2]

1 Corporate Technology, Siemens AG, Günther-Scharowsky-Str. 1, Erlangen, Germany

2 Faculty of Applied Computer Science, University of Augsburg, Augsburg, 86159, Germany

**ABSTRACT**

For a wider utilization of model-based control e.g. for energy management systems, their installation and maintenance costs must be reduced. A possible solution is the automated identification of the different asset model parameters. However, robust model validation methods are necessary in order to guarantee an adequate performance in practice without additional manual review. This paper presents a model validation method for energy conversion units based on an uncertainty and consistency analysis of the extrapolated energy efficiency ratio (EER). First, the approach is described in detail. Afterwards, the advantages of the concept over validation methods based on model accuracy are illustrated with a case study of a compression chiller. Only the presented approach ensured a robust validation of the chiller models.

**Keywords:** model predictive control, automated parameter identification, energy management system, compression chiller, confidence interval

## 1. INTRODUCTION

### *1.1 Motivation for an automated parameter identification in model-based control*

Many studies have proved that energy management systems based on model predictive control (MPC) can reduce energy consumption of building complexes [1]. As with all model-based controls, the control quality is significantly dependent on the accuracy of the different asset models. A uniform modelling of these assets with an a priori determined parameter set is therefore not recommended. Especially since their operating behaviour also varies, e.g. due to different configurations and installation conditions. At the same time, manual identification of the parameters increases both the costs and the effort of commissioning substantially. Energy management systems based on an MPC are as a result only primary used in scientific and feasibility studies [2].

An automated identification of the different model parameters could reduce these costs considerable. However, in practice, robust model validation methods are required in order to ensure an error-free operation of the MPC without additional manual review.

### *1.2 Drawbacks of model validation approaches based on model accuracy*

The operating points of MPCs vary over time and can be frequently at the borders of the operating range (OPR). Identified models should thus be valid for the entire OPR to avoid detrimental control performance. Conventional model validation approaches based on model accuracy assessment (e.g. root mean square error (RMSE)) are usually only robust if the training data covers all operating points. This applies especially to black-box models [3]. Consequently, the training data for these validation methods should consist of the complete OPR. However, this is in many cases either impractical or very time-consuming. Particularly in brownfield applications, the operating points can only be varied to a limited extent during operation. Furthermore, MPCs are commonly supervisory controller and therefore don't have direct control access to all exogenous input variables of the models. The control variable for compression chillers are for instance frequently only on/off. In this case, the capacity of the chiller is solely adjusted by the local controller to meet the cooling demand. Thus, a variety of different load profiles would be necessary in order to obtain training data for the complete cooling capacity range. Additionally, an automated determining and distinguishing of explicit operating points is challenging with increased model complexity and number of exogenous input variables due to e.g. multicollinearity.

Hence, a stochastic validation approach which quantifies the information content of the training data for a given model structure including the correlation and variation of the exogenous variables is required.

This paper presents a model validation method for energy conversion units using compression chillers as an example. The approach is based on an uncertainty and consistency analysis of the energy efficiency ratio (EER) extrapolated over the complete OPR. The subsequent section outlines how the uncertainty of the $\mathrm{EER}$ can be determined, and which performance criteria can be applied to it. Afterwards the approach is validated in section 3 with a case study of a compression chiller.

## 2. METHODOLOGY

### 2.1 Definition of an energy conversion unit

An energy conversion unit transforms one or multiple energy flows $\boldsymbol{p}_{\mathbf{in}} \in \mathbb{R}^l$ into one or multiple output energy flows $\boldsymbol{p}_{\mathbf{out}} \in \mathbb{R}^m$. The respective EERs

$$\mathrm{EER}_{a,b} = \frac{p_{\mathrm{out},a}}{p_{\mathrm{in},b}},\ \forall a \in [0..m], b \in [0..l] \tag{1}$$

can be modelled as quasi-steady state functions $\mathbf{EER}(\boldsymbol{\theta}, \boldsymbol{v}, \boldsymbol{p}_{\mathbf{in,out}}) \in \mathbb{R}^{l\cdot m}$ of model parameters $\boldsymbol{\theta}$, exogenous inputs $\boldsymbol{v}$ (e.g. ambient temperature) and either $\boldsymbol{p}_{\mathbf{in}}$ or $\boldsymbol{p}_{\mathbf{out}}$. Practical examples of energy conversion units include chillers, heat pumps, diesel generators, and combined heat and power units.

The following validation approach applies to one steady state function $\mathrm{EER}_{a,b}$ respectively and should therefore be carried out individually for all of them.

### 2.2 Identification of model parameters

To estimate $\boldsymbol{\theta} \in \mathbb{R}^k$ the model is usually transformed into a linear-in-parameter formulation

$$\boldsymbol{y} = \boldsymbol{X}\widehat{\boldsymbol{\theta}} + \hat{\boldsymbol{\epsilon}}, \tag{2}$$

where the vector of dependent variables $\boldsymbol{y} \in \mathbb{R}^n$ as well as the matrix of independent variables $\boldsymbol{X} \in \mathbb{R}^{n\times k}$ are functions of $\boldsymbol{v}$ and $\boldsymbol{p}_{\mathbf{in,out}}$. Subsequently, by using the ordinary least square (OLS) method, which minimizes the sum of estimated squared residuals $\hat{\boldsymbol{\epsilon}} \in \mathbb{R}^n$, $\widehat{\boldsymbol{\theta}}$ can be calculated as follows:

$$\widehat{\boldsymbol{\theta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \tag{3}$$

### 2.3 Estimation of the parameter uncertainty

To determine the respective $\mathrm{EER}$ uncertainty, the covariance matrix of $\widehat{\boldsymbol{\theta}}$

$$\begin{aligned}\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} &= \mathrm{E}\left(\left[\widehat{\boldsymbol{\theta}} - \mathrm{E}(\widehat{\boldsymbol{\theta}})\right]\left[\widehat{\boldsymbol{\theta}} - \mathrm{E}(\widehat{\boldsymbol{\theta}})\right]'\right) \\ &= (\boldsymbol{X}'\boldsymbol{X})^{-1} \cdot \boldsymbol{X}'\mathrm{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}')\boldsymbol{X} \cdot (\boldsymbol{X}'\boldsymbol{X})^{-1}\end{aligned} \tag{4}$$

must be calculated beforehand. The difficulty thereby is the estimation of the covariance matrix of the residuals of the parent population $\mathrm{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}')$. OLS assumes that the residuals have a constant variance (homoscedasticity) and no autocorrelation. Under these circumstances, (2) simplifies to:

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} = \sigma_{\boldsymbol{\epsilon}}(\boldsymbol{X}'\boldsymbol{X})^{-1}, \tag{5}$$

where $\sigma_{\boldsymbol{\epsilon}}$ is the standard deviation of the residual of the parent population. However, due to model simplifications (e.g. quasi-steady state and linear model structure) these conditions often do not exist in practice. In these cases, OLS is still an unbiased and consistent estimator but no longer efficient. As a result, the calculated $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$ in (5) frequently understates the true standard deviations of $\widehat{\boldsymbol{\theta}}$ and is no longer reliable. In order to obtain a better estimate of the true uncertainty, so-called heteroskedasticity and autocorrelation consistent (HAC) standard errors can be used [3]. The most common approach is the Newey-West estimator:

$$\boldsymbol{X}'\mathrm{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}')\boldsymbol{X} = \textstyle\sum_{i,j=1}^{n} w_{|i-j|} \cdot (x_i\hat{\epsilon}_i)(x_j\hat{\epsilon}_j)'. \tag{6}$$

Here, $n$ is the number of measurement data set, $x_i$ is the independent variables of the data set $i$ and

$$w_i = 1 - \frac{i}{l+1} \tag{7}$$

are linear decaying weights with the bandwidth $l$ used to compensate the influence of an autocorrelation over $l$ lags [4]. Thus, all weights beyond $l$ are set to zero.

### 2.4 Determining the EER uncertainty

The standard deviation of $\mathrm{EER}$ can be derived by means of error propagation from $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$ as follows:

$$\hat{\sigma}_{\mathrm{EER}}(\widehat{\boldsymbol{\theta}}, \boldsymbol{v}, \boldsymbol{p}_{\mathbf{in,out}}, \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}) = \sqrt{\left.\frac{\partial \mathrm{EER}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} \cdot \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} \cdot \left.\frac{\partial \mathrm{EER}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}'}. \tag{8}$$

The uncertainty of $\mathrm{EER}$ with a $(1-\alpha)$ confidence level is obtained on the assumption of a t-distribution:

$$\begin{aligned}U_{\mathrm{EER}}(\widehat{\boldsymbol{\theta}}, \boldsymbol{v}, \boldsymbol{p}_{\mathbf{in,out}}, \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}) &= \mathrm{EER} \pm t_{1-\frac{\alpha}{2}} \cdot \hat{\sigma}_{\mathrm{EER}} \\ &= \mathrm{EER} \pm \frac{\delta}{2}.\end{aligned} \tag{9}$$

*2.5 Validation criteria based on EER uncertainty*

To assess the model validity, the determined $\mathrm{EER}(\widehat{\boldsymbol{\theta}}, \boldsymbol{v}, \boldsymbol{p}_{\mathrm{in,out}})$ and $U_{\mathrm{EER}}(\widehat{\boldsymbol{\theta}}, \boldsymbol{v}, \boldsymbol{p}_{\mathrm{in,out}}, \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}})$ are first extrapolated over the complete OPR of the energy conversion unit. Hence, a basic prerequisite of the used model structure is the ability to adequately represent the behaviour of the energy conversion unit over this range. Afterwards, the following two criteria are examined:

The first criterion evaluates $U_{\mathrm{EER}}$ regarding consistency based on logically defined limits (e.g. Carnot efficiency)

$$\mathrm{EER}_{\min}(\boldsymbol{v}, \dot{Q}) \leq U_{\mathrm{EER}} \leq \mathrm{EER}_{\max}(\boldsymbol{v}, \dot{Q}). \tag{10}$$

This is particularly advantageous for black box models, as their parameters do not enable consistency checks.

The second criterion reviews whether the relative uncertainty does not exceed a selected limit (e.g. 10 %)

$$\frac{\delta}{2 \cdot \mathrm{EER}} \leq U_{\max}. \tag{11}$$

If both criteria are met, the model will be assumed as valid. A two-dimensional example of the presented approach with one exogenous input signal is illustrated in Fig. 1.

## 3. CASE STUDY

### *3.1 Experimental setup*

#### *3.1.1 Description of used data sets and models*

The validation method presented in the previous section was evaluated using 22 months of measured operational data of an air-cooled compression chiller installed in a business facility in Berlin Adlershof. During this period the supervisory controller was still rule based.

The EER of a chiller, known as the coefficient of performance (COP), is the ratio of evaporator cooling capacity $\dot{Q}$ to the electrical input power of the compressor $P$. In order to describe the compression chiller, the universal Gordon-Ng model (GNU) (12), its simplified empirical version (GNS) (13) and a self-derived multivariate polynomial model (MP) (14) were selected [5]. While the GNU model is a grey-box model with physically interpretable parameters, the other two are black-box models and linear regarding $P$ and $\dot{Q}$, which are often optimization variables of an MPC.

$$\frac{T_{\mathrm{e}}\left(1+\frac{1}{\mathrm{COP}}\right)}{T_{\mathrm{c}}} - 1 = \theta_1 \frac{T_{\mathrm{e}}}{\dot{Q}} + \theta_2 \frac{T_{\mathrm{c}} - T_{\mathrm{e}}}{T_{\mathrm{c}}\dot{Q}} + \theta_3 \frac{\dot{Q}\left(1+\frac{1}{\mathrm{COP}}\right)}{T_{\mathrm{c}}} \tag{12}$$

$$\frac{1}{\mathrm{COP}} = -1 + \frac{T_{\mathrm{c}}}{T_{\mathrm{e}}} + \frac{1}{\dot{Q}}\left(-\theta_1 + \theta_2 T_{\mathrm{c}} - \theta_3 \frac{T_{\mathrm{c}}}{T_{\mathrm{e}}}\right) \tag{13}$$

$$\frac{1}{\mathrm{COP}} = \theta_1 + \theta_2 T_{\mathrm{e}} + \theta_3 T_{\mathrm{c}} + \theta_4 \frac{T_{\mathrm{c}}}{T_{\mathrm{e}}} + \theta_5 T_{\mathrm{e}}^2 + \theta_6 T_{\mathrm{c}}^2 + \frac{1}{\dot{Q}}\left(\theta_7 + \theta_8 T_{\mathrm{c}} + \theta_9 \frac{T_{\mathrm{c}}}{T_{\mathrm{e}}}\right) \tag{14}$$

$T_{\mathrm{e}}$ is here the water outlet temperature of the evaporator and $T_{\mathrm{c}}$ the inlet temperature of the condenser. In this case study, $T_{\mathrm{c}}$ is assumed to equal the ambient temperature since the chiller is air-cooled.

During data preparation, the sections with sensor errors and the days with the chiller being switched off were first removed from the data. Afterwards a finite impulse response low pass filter was used to obtain a stationary behaviour.

#### *3.1.2 Test procedure*

The following test procedure was used to assess the robustness of the presented validation method:

Step 1: Identification of the GNU, GNS and MP model with an increasing set of training data, until the validation criterion is fulfilled

Step 2: Assessment of the model quality using the remaining measurement as test data and the coefficient of variation (CV) of the identified $\widehat{\mathrm{COP}}$:

$$\mathrm{CV} = \frac{1}{\overline{\mathrm{COP}}}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\mathrm{COP}}_i - \mathrm{COP}_i\right)^2} = \frac{\mathrm{RMSE}_{\mathrm{COP}}}{\overline{\mathrm{COP}}} \tag{15}$$

In order to obtain a variety of test instances, Steps 1 and 2 were repeated continuously with the start date increased by one (see Fig. 2). The used parameters of the

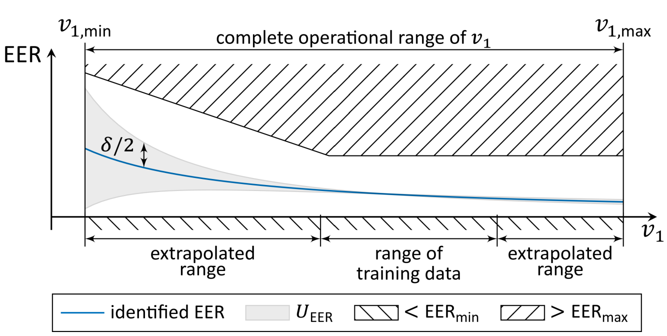


Fig. 1: Example of the two validation criteria with one exogenous input $v_1$. If $U_{\mathrm{EER}}$ is within logically defined limits $[\mathrm{EER}_{\min}, \mathrm{EER}_{\max}]$ and the relativ uncertainty $\left(\frac{\delta}{2 \cdot \mathrm{EER}}\right)$ is smaller than a select limit, the model is assumed as valid.

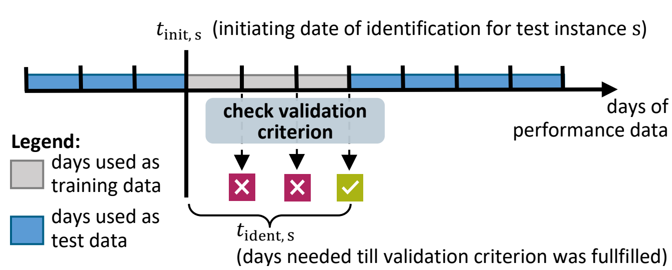


Fig. 2: Test procedure for each validation method. The initiating date of the identification increases for each test instance $s \in [1,222]$ by one day.

Table 1 Experimental Setup

| | | |
|---|---|---|
| Chiller Type | | Trane RTAC 170 |
| days of performance data | | 222 |
| OPR: | $\dot{Q}$ (kW) | [80, 535] |
| | $T_c$ (°C) | [-5, 35] |
| | $T_e$ (°C) | [-6, -4] |
| $COP_{min}$/ $COP_{max}$ | | 0/6 |
| $U_{max}$ | | 15 % |
| $(1-\alpha)$ | | 95 % |

validation criterion are listed in detail in Table 1. The limit for the validation criterion $COP_{max}$ was defined as 200 % of the nominal COP. The bandwidth of the estimator was calculated according to the recommendation of Newey and West [3].

To assess the presented method further, the test procedure was also applied to two conventional validation methods based on model accuracy. These methods assume the model as valid if:

(a) CV of the COP < 5 %
(b) average CV of 2-fold cross-validation < 5 %

### 3.2 *Results and discussion*

As shown in Figure 3, only the presented validation method based on a consistency and uncertainty analysis ensured a robust model validation.

In several test instances, the conventional validation methods have classified the model prematurely and incorrectly as valid. This led especially for the GNS and MP model to outliers with a high model deviation. However, even the grey-box GNU model could not be validated robustly with the methods based on model accuracy. Early use of these faulty models in an MPC application significantly deteriorates the control performance and can even cause errors.

Consequently, the additional amount of used training data by the validation approach based on uncertainty was needed in order to obtain a better model for the OPR. To shorten the identification time in practice, one can also use a limited OPR based on the predicted value ranges of the exogenous inputs of the near future. However, in this case, continuous identification and validation is inevitable while adjusting the respective current OPR.

The presented validation approach in this paper ensured a robust validation without additional manual review. In the future it will also be evaluated for other energy conversion units and tested in combination with an MPC.

## ACKNOWLEDGEMENT

The authors would like to thank the German Federal Ministry for Economic Affairs and Energy (BMWi) supporting the project "Energienetz Berlin Adlershof - Monitoring und Optimierung" (no. 03ET1549C).

## REFERENCE

[1] A. Afram and F. Janabi-Sharifi. Gray-box modeling and validation of residential HVAC system for control system design. Appl. Energy, vol. 137, pp. 134–150, 2015.

[2] J. T. Wen and M. Sandipan. Intelligent Building Control Systems - A Survey of Modern Building Control and Sensing Strategies. Springer International Publishing AG, 2018.

[3] W. J. Den Haan and A. T. Levin. A Practitioner' s Guide to Robust Covariance Matrix Estimation. vol. 15, no. 1988, 1997.

[4] A. Zeileis. Econometric Computing with HC and HAC Covariance Matrix Estimators. J. Stat. Softw., vol. 11, no. 10, 2004.

[5] J. M. GORDON and K. C. NG. Cool thermodynamics - the engineering and physics of predictiv, diagnostic and optimization methods for cooling systems. 1st ed. Cambridge: Cambridge International Science Publishing, 2001.

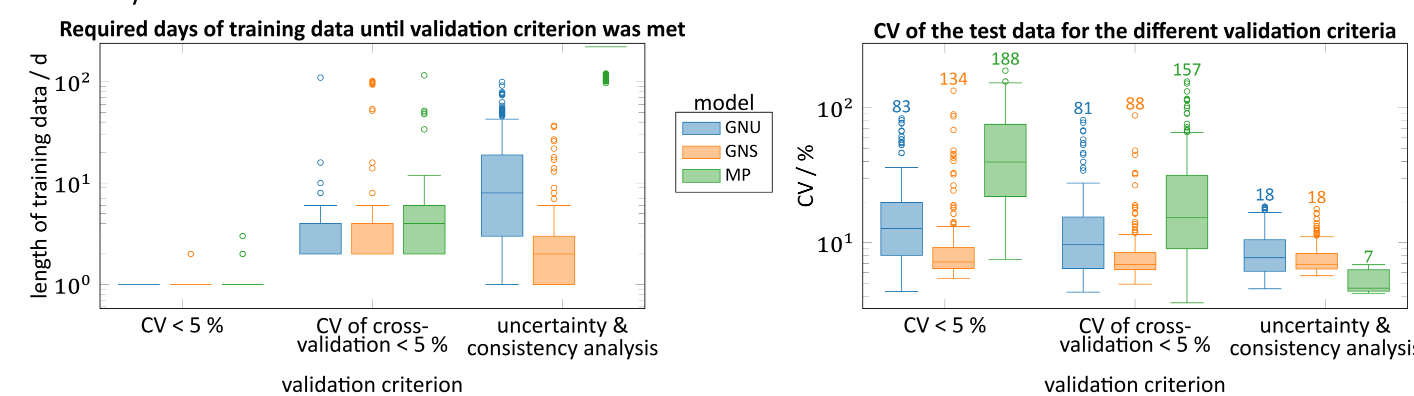


Fig 3 Required days of training data until the validation criteria were fullfilled and comparison of the CV value for the as valid assumed models for test data. Only the presented validation approach based on uncertainty and consistency ensured a robust validation, since there are no significant outliers of the CV value.